JOURNAL OF FOOD COMPOSITION AND ANALYSIS (2002) **15,** 435–451
doi:10.1006/jfca.2002.1089
Available online at http://www.idealibrary.com on IDEAL®

STUDY REVIEW

The European Nutrient Database (ENDB) for Nutritional Epidemiology

U. R. Charrondiere* J. Vignat*, A. Møller†, J. Ireland‡, W. Becker§, S. Church¶, A. Farran‖, J. Holden**, C. Klemm††, A. Linardou‡‡, D. Mueller§§, S. Salvini¶¶, L. Serra-Majem‖‖, G. Skeie***, W. van Staveren†††, I. Unwin‡‡‡, S. Westenbrink§§§, N. Slimani*, and E. Riboli*,[1]

*International Agency for Research on Cancer (IARC) WHO, Lyon, France; †Institute of Food Safety and Nutrition, Danish Veterinary and Food Administration, Søborg, Denmark; ‡Informatic Center for Food Quality (CIQUAL), French Food Safety Agency (AFSSA), Maisons-Alfort, France; §Swedish National Food Administration, Uppsala, Sweden; ¶Nutrition Division, Food Standards Agency, London, U.K.; ‖Center d'Ensenyament Superior de Nutrició i Dietètica, Universitat de Barcelona (CESNID-UB), Santa Coloma de Gramanet, Spain; **USDA, Nutrient Data Laboratory, Beltsville, U.S.A.; ††Federal Institute for Health Protection of Consumers and Veterinary Medicine (BGVV), Berlin, Germany; ‡‡Department of Hygiene and Epidemiology, University of Athens Medical School, Goudi-Athens, Greece; §§Procter & Gamble, Schwalbach, Germany; ¶¶Epidemiology Unit, Centro per lo Studio e la Prevenzione Oncologica, Firenze, Italy; ‖‖Center for Community Nutrition Research, University of Barcelona, Barcelona, Spain; ***Institute of Community Medicine, University of Tromsø, Tromsø, Norway; †††Department Human Nutrition and Epidemiology, Wageningen University, Wageningen, The Netherlands; ‡‡‡Food Information Consultancy, Cambridge, U.K.; and §§§TNO Nutrition and Food Research Institute, NEVO Foundation, Zeist, The Netherlands

Received December 21, 2001, and in revised form May 24, 2002

Food composition databases (FCDB), as well as standardized calculation procedures are required for international studies on nutrition and disease to calculate nutrient intakes across countries. Comparisons of national FCDBs have shown that major improvements are needed in standardization and documentation at the food and nutrient levels to minimize systematic and random errors in nutrient intake estimations. The International Agency for Research on Cancer (IARC), together with national FCDB compilers, researchers in international studies (EURALIM, SENECA) and industry, is currently developing a standardized and critically assessed nutrient database for the 10 countries involved in the European Prospective Investigation into Cancer and Nutrition (EPIC): Denmark, France, Germany, Greece, Great Britain, Italy, The Netherlands, Norway, Spain and Sweden. It will be compiled using the general concept for a standardized FCDB, food classification and description, and calculation procedures developed for EPIC. National compilers will provide and document a subset of their nutrient data and some will evaluate them. Updated 'Food Table Input' (FTI) software will be used to evaluate and compile the data. The European Nutrient Database (ENDB) will contain values for approximately 100 nutrients for 1000 foods per country, which is mainly derived from EPIC consumption data. In the future, this database could be extended to

[1]To whom correspondence and reprint requests should be addressed. Tel.: +33-4-72-73-84-11; Fax: +33-4-72-73-83-61; E-mail: riboli@iarc.fr

0889-1575/02/040435 + 17 $35.00/0

include more foods, components and countries. Additionally, methodological issues should be addressed elsewhere and awareness of the need for standardizing FCDBs and their procedures has to be increased among users and funding agencies.

Key Words: nutrient intake; standardization; food composition table; nutrient database; EPIC.

INTRODUCTION

In Europe, most countries have national food composition tables, which are compiled according to country-specific procedures and traditions. Their primary objective is to provide comparable nutrient composition data over time at the national level. They were not necessarily conceived to provide internationally comparable data. Since 1982, several international initiatives have been undertaken to harmonize procedures for better data comparability and interchange. These initiatives are International Network of Food Data Systems (INFOODS) with their regional branches, now coordinated by Food and Agriculture Organization (FAO) in Rome. In Europe, EUROFOODS projects were funded as concerted action projects from the Commission of the European Community under the 'Eurofoods-Enfant project within FLAIR' (Food-linked Agro-Industrial Research) and 'COST ACTION 99' (European co-operation in the Field of Scientific and Technical Research). A subregional branch, NORFOODS, is working on harmonizing data within the North European countries. In addition, international and national institutes have increased their international collaboration. These initiatives have advanced the harmonization in several fields: analytical methods, nutrient nomenclature, definitions and mode of expression, food nomenclature, description, terminology and classification, and database management and interchange. INFOODS provided guidelines on the organization and content of food composition tables and databases, methods for analysing foods and compiling those tables, and procedures for accurate international interchange of data (Lupien, 1995), and developed the INFOODS tagnames (Klensin, 1992). They also stimulated the development of intake estimation software as well as the publication of regional food composition tables through ASEANFOODS (1996), LATINFOODS (2000) and CEECFOODS (in prep.). Major deliverables from Eurofoods-Enfant and COST Action 99 are the Eurofoods-Enfant report on nutrient losses and gains in the preparation of foods (Bergström, 1994), a set of recommendations defining basic data structures and format for data interchange (Schlotke *et al.*, 2000) and the book by Greenfield and Southgate (1992): 'Production, Management and Use of Food Composition Data', which was produced in co-operation with INFOODS.

However, due to lack of funding and a driving force, no standardized FCDB has been developed in Europe so far, even though its creation was planned within the COST ACTION 99 (EUROFOODS) project. Only when EPIC, with food consumption data from 10 European countries (Riboli and Kaaks, 1997), clearly stated that a standardized FCDB is needed to derive comparable nutrient values across countries (Slimani *et al.*, 1995) because nutrient values in national FCDBs are not truly comparable (Deharveng *et al.*, 1999), was the development of a standardized FCDB started (Slimani *et al.*, 2000a).

The lack of a standardized European FCDB represents a handicap for international nutritional epidemiology in the investigation of the relationship between nutrition and health, and the comparison and evaluation of nutrient intakes. Even if specific nutrient intakes are not investigated, energy has to be calculated from macronutrients because it is widely used as an adjustment factor in food intake estimations (Willett, 1998) and nutrient composition of foods has to be known to relate biomarker measurements to food intake. As nutrient exposures ultimately depend on the quality and use of food composition tables (Bingham, 1987; Cameron and Van Staveren, 1988; Haraldsdottir, 1993) as well as the intake estimation system in use (NORFOODS, 2002), measurement errors in nutrient intake between population groups may be one of the reasons for many contradictory results on the relation between diet and chronic disease (Riboli, 1989). Systematic and random errors in food composition tables and consequently in nutrient intake estimations may bias the relationship between diet and disease, e.g., a shift in relative risk of disease and in the ranking of subjects in terms of disease risk. These errors in nutrient values may then result in a false conclusion concerning the relationship between diet and disease, or incorrect nutrient estimations leading to erroneous public health measurements. However, researchers in nutritional epidemiology, politicians and funding agencies are not fully conscious of the actual problems concerning the incompatibility and imprecision of compositional data and nutrient intake measurements across countries.

In this paper, concerns regarding the international use of existing food composition data are presented as well as how the ENDB project intends to overcome the different aspects of converting food intake to comparable nutrient intake across countries.

THE VARIATION OF NUTRIENT INTAKE ESTIMATIONS DUE TO DIFFERENCES IN NATIONAL FOOD COMPOSITION TABLES

The comparison between food composition tables in nine European countries showed that nutrients differ in definition, analytical methods, units and mode of expression which could potentially lead to different nutrient values between tables (Deharveng *et al*., 1999). In addition, a wide variation is observed with respect to, e.g., number of foods, language, food classification and description, and if mainly raw foods or also cooked foods with very different nutrient composition are included. These differences may have an impact on the precision of nutrient intake estimations (Slimani *et al*., 2000a) and make international comparisons difficult and imprecise (EFCOSUM group, 2001).

A series of comparisons between nutrient calculation softwares and computerized food composition databases from the same country showed conclusively that intakes for certain nutrients differ significantly, which are accentuated through different application of the compositional data in nutrient intake calculations by users (Adelman *et al*., 1983; Hoover, 1983; Shanklin *et al*., 1985; Eck *et al*., 1988; Kohlmeier, 1991; Nieman *et al*., 1992; Guilland *et al*., 1993; Lee *et al*., 1995; Granado *et al*., 1997). More recently, comparison of different national FCDBs arrived at similar results, probably detecting even bigger differences, up to 20–45% in mean intakes of individual nutrients for the same diet (see Table 1). The variation in nutrient intake can be attributed to real differences in the composition of foods between countries but the same artificial differences as mentioned above seem to play an important role (Vandenlangenberg *et al*., 1996; Charrondiere *et al*., 2001a, b; NORFOODS, 2002). These big differences were not found for macronutrient intakes in the SENECA study when comparing nutrient intakes calculated by three national tables and to the Dutch table. Mean intakes differed less than 10%, which was

TABLE 1

Mean nutrient intake of 1314 British EPIC subjects by applying the British, Danish and French FCDBs for five food groups such as potatoes, vegetables, pulses, fruits, and cereals

	British FCT	French FCT	Danish FCT
Energy (kJ)	3063*	3206*	3736*
Protein (g)	22.5	21.6*	22.7
Fat (g)	12.1*	11.7*	13.0*
Carbohydrates (g)	139.6	139.3	168.0*
Fibre (g)	14.4* (Englyst) 20.5* (Southgate)	20.0 (AOAC)	20.0 (AOAC)
Vitamin C (mg)	75.4*	86.2*	109.0*
Potassium (mg)	1521*	1576*	1452*

*Significant difference with both other tables ($P < 0.0001$).
Source: Charrondiere *et al.* (2001b).

considered acceptable since a certain proportion of the variation is due to true differences in foods. However, differences in conversion factors were corrected beforehand and nutrient values were calculated based on ingredients (Moreiras *et al.*, 1991).

It can therefore be concluded that the differences observed in nutrient values between the national FCDBs reflect not only the actual differences between foods due to growing, harvesting, storing, and processing conditions but that nutrient values also differ artificially between tables. These differences are further accentuated due to missing foods or values, which force the users of FCDBs to develop their own procedures for calculations, for treating missing values and food matching. A standardized European food composition database is thus needed for multi-national studies on nutritional epidemiology in Europe to calculate comparable nutrient intakes. Based on a theoretical concept developed for such a database (Slimani *et al.*, 2000a), fruitful collaboration at the European level has created the basis of the ENDB project.

OBJECTIVE OF THE ENDB

The objective of this project was to develop a standardized and critically assessed European nutrient database (ENDB) for the 10 countries involved in EPIC (Denmark, France, Germany, Greece, Great Britain, Italy, The Netherlands, Norway, Spain and Sweden). The ENDB will be a database comprised of 11 national data sets based on the compositional data from the national FCDBs. The USDA database has also been included in the project as it is frequently used as a source for missing data. Nutrient values will be compiled using the EPIC food classification and description for at least 1000 consumed foods per country, derived mainly from EPIC and completed by foods from other international studies (EURALIM and SENECA). All nutrient values will be documented, standardized and evaluated according to common rules without missing values for important nutrients, e.g., macronutrients and most vitamins and minerals (but not their contributing components). It is planned that the ENDB will be completed with first priority nutrients, mainly macronutrients, by 2002 and with second priority nutrients, namely vitamins and minerals, by 2003.

TABLE 2

Differences between national FCDBs and how ENDB plans to overcome them

	Differences in National FCDBs	ENDB objectives
At food level	Between 60 and 11,000 foods, some in local language only	At least 1000 foods per country in local language, English and taxonomic names
	Incomparable food classification and description	Common food classification and description
	Some with mainly raw foods, others with foods as consumed	Foods as consumed
	Coverage of frequently consumed foods sometimes insufficient, especially for meat and manufactured products	Sufficient coverage of frequently consumed foods. Nutrient values also from other sources, e.g., food industry, other databases
At nutrient level	Different definitions and analytical methods	Different definitions and analytical methods to be *separated* on nutrient level
	Different modes of expression and units	Standardized modes of expression and units
	Missing nutrient values and outdated values	Calculation or estimation of missing nutrient values and replacement of outdated values by data from other sources
Documentation of nutrient values	Lacking in most FCDBs for:	Documentation provided as far as possible:
	Definitions, analytical methods, source of value at nutrient level (e.g., if analysed, calculated, copied)	At the nutrient level definitions, analytical methods, and sources to evaluate quality and comparability of values
	Food sampling methods usually not available	Food sampling methods if existing
	Algorithms, coefficients or recipes used to calculate missing values are lacking in most FCDBs	Algorithms, coefficients or recipes used to calculate missing values
At compilation level	Different treatment of missing values	No missing values, to avoid underestimation of nutrient intake
	Different calculation procedures, e.g., recipe/ algorithm calculations, retention factors	Common calculation procedures, algorithms and coefficients

The ENDB project aims to overcome at the international level the differences encountered between the national FCDBs, which should result in more comparable nutrient intake data with minimal random and systematic errors. These differences can be grouped into four major categories as shown in Table 2.

PARTNERS OF THE ENDB PROJECT

The ENDB project, which is coordinated by IARC, involves 11 compilers of national FCDBs, three international studies (EPIC, SENECA and EURALIM), one information specialist and one representative of the food industry (see Appendix 1). In the future, it is expected that more partners from the food industry will join the project and provide compositional data.

NUTRIENTS IN THE ENDB

The ENDB will concentrate on a choice of nutrients, which have been selected for their importance in nutritional epidemiology (Slimani *et al.*, 2000a) and then subdivided into first and second priority nutrients (see Appendix 2). First priority nutrients, mainly macronutrients, are needed to calculate energy intakes. They will also be used to investigate the first set of cancer hypotheses on macronutrients and dietary fibre. The second priority nutrients are mainly micronutrients and will be important to relate biomarkers and to investigate further cancer hypotheses.

For each of these selected nutrients, mode of expression and units will be harmonized and if necessary, nutrients will be reported separately according to their definition and analytical method. This distinction will be necessary for, e.g., energy, carbohydrates, dietary fibre, Vitamin D or folate as their values depend heavily on the analytical method and definition used. In some cases, conversion factors exist to adopt the value from one method to another like the regression equation of Mongeau and Brassard (1989) for dietary fibre. The values of nutrients like energy or protein will be calculated according to given definitions and conversion factors, i.e., their values will not be imported directly from national FCDBs, in order to avoid systematic errors.

The ENDB aims to fill in all values for important nutrients to avoid an underestimation of nutrient intakes due to missing values, which is a source of systematic error. Charrondiere *et al.* (2001a) have demonstrated for dietary fibre that treating missing values as zero can result in a 20% lower mean nutrient intake compared to computing the value from a similar food. The degree of under-estimation depends on the proportion of missing values and on the frequency of consumption of foods with incomplete data (Kohlmeier, 1991).

FOODS IN THE ENDB

The majority of foods in the ENDB will derive from the EPIC main dietary questionnaires and 24-h-recall records. The 24-h-recall interviews were carried out with EPIC-SOFT, a software specially designed within the EPIC project to collect 24-h-recall interviews, standardized between countries in terms of food nomenclature, classification, description and quantification (Slimani *et al.*, 1999, 2000b). In EPIC-SOFT, the ingredients of mixed recipes were treated as foods, which makes the food intakes more comparable among countries and is easier for the attribution of nutrient values. This means that the ENDB will not contain mixed recipes or dishes. Depending on the country, between 5000 and 12000 foods and ingredients were reported, including reported brand and product names, from which important foods will be selected (Charrondiere *et al.*, 2000). Further foods will be added such as ingredients of recipes to calculate nutrient values of complex foods such as cakes, foods with a specific hypothesis on cancer etiology, and important foods from EURALIM and SENECA, the other international studies involved in the project. This process will result in about 1000 foods per country as entries to the ENDB.

In addition, foods will be included into the ENDB for methodological reasons, to obtain the documentation from the compiler of the original nutrient values, which are copied from one table into another. The countries most concerned are the U.S.A. and the U.K. as their databases are frequently used as reference.

It was decided to attribute to each national food as much as possible country-specific nutrient values, i.e., to keep their nutrient values of the national food

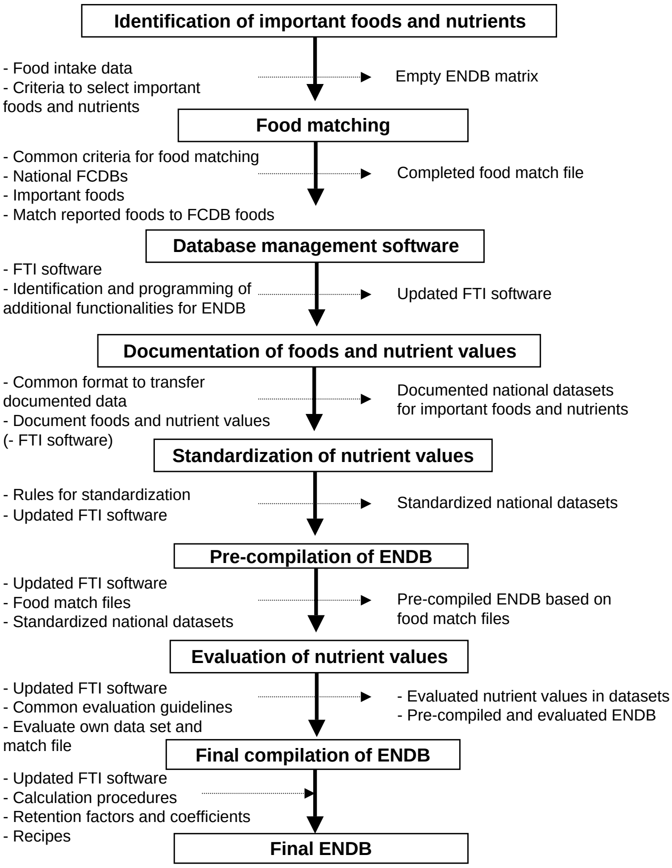

FIGURE 1. Flow of ENDB project.

composition table if proven adequate. Therefore, if a food with the same name (for example, 'white bread') was reported in several countries, it will be listed with different country codes. This approach is thought to be an advantage because country-specific food varieties and conditions can be taken into account, instead of recording only one food to be applicable for all European countries (see Fig. 1).

FOOD MATCHING IN THE ENDB

The food matching is an important step in the quality of calculations of nutrient intake estimations. The process of linking foods, for which nutrient intakes are to be estimated, to foods present in food composition tables influences the validity of nutrient intake estimations. If the exact corresponding food is not found in the table, a food match has to be done with a similar food, which most probably has a different nutrient composition, or its nutrient values have to be calculated. Both approaches lead to an approximation of the real nutrient composition. It can be assumed that measurement errors in nutrient intake are high when the percentage of exact food matches with foods in the table is low. Consequently, the quality of nutrient intake estimations is influenced by the degree that foods are covered by the national tables and the knowledge of the person carrying out the food matching about the foods available in the country and in the FCDB. Therefore, in the ENDB project, an approach was chosen that allows a standardized and high-quality food match. The food matching will be carried out according to common rules (Internal report, 2001) and supervised by the national compiler who knows most about local foods. In addition, other national data sets are available as sources of food matches. In a first phase, as many as possible of the ENDB foods are matched by the compiler and/or by the local EPIC centre with foods in the national FCDB. A food can be matched to an equivalent or similar food, or to a food used to calculate the missing values. For some foods, no matches will be indicated either because the nutrient values will be calculated based on recipes or because no food match is possible with any food from the national table. During the data evaluation, compilers will complete missing food matches and replace food matches with 'similar' foods with well corresponding foods from other national data sets. In this way, the quality of food matches in ENDB will be increased as compared to food matches with the sometimes limited number of foods covered in the national FCDB.

The first phase of the food matching allows the identification of foods from the national FCDB, for which nutrient values will have to be documented for this project. At a later stage, the corrected food matches will be used for the automatic compilation of the nutrient values of the ENDB database.

DOCUMENTATION OF NUTRIENT VALUES AND FOODS DERIVED FROM NATIONAL FCDBs

Only a full documentation of foods and nutrient values in a database can lead to well-founded judgements on the appropriateness of values and the justification of discrepancies between values (Southgate, 1993). Therefore, when possible, in the ENDB the original sources of nutrient values will be documented, as well as the definition and analytical methods used, the number of samples and the range of values.

The participants in the ENDB project have agreed to document a subset of foods and their nutrient values from their national FCDB. In order to carry out the documentation in a standardized way and format, the 'Guideline notes for preparing and exporting food composition data' have been developed (Vignat *et al.*, 2001). They are based on the general recommendations laid down in the much broader and more complete 'EUROFOODS Recommendations for Food Database Management and Data Interchange' (Schlotke *et al.*, 2000). The ENDB has adapted and expanded the EUROFOODS component tagnames and the list of analytical methods. Data of manufactured foods, apart from the source of the company, will probably not be documented.

Additional information to be made available if existing:

- the algorithms, together with the weight yield factors, which transform the raw food weight into the cooked food weight,
- the density or specific gravity coefficients,
- the coefficients on the amount of fat absorbed into the food during cooking,
- the nutrient retention factors for the different cooking methods, and

- all recipes with their ingredients and proportions used to calculate missing nutrient values for the complex foods (e.g., cakes, sauces, soups) that were not broken down into ingredients during the data collection.

DATABASE MANAGEMENT SOFTWARE

Data preparation for the ENDB includes major data handling steps that require the support of data management software. These steps are documentation, standardization and evaluation of national data sets and compilation of the ENDB from the evaluated data sets. Instead of developing a completely new database management system for this project, it was decided to adapt the existing software 'Food Table Input' (FTI) to meet the requirements for the ENDB project. The FTI software is already capable of storing and displaying data from several national databases, with fields for handling metadata such as food names in multiple languages, food description, analytical methods and bibliographic references. It also complies with both the EUROFOODS "Recommendations for Data Interchange and Management" and the ENDB "Guideline notes for preparing and exporting food composition data".

The further development of the FTI will include specific functions for the ENDB project such as the preliminary compiled ENDB based on the food match file with data from national databases, specific functionalities concerning the EPIC study as well as control functions and the ENDB calculation functions. With the updated FTI software, every compiler can conveniently evaluate data by comparing documented data from a range of data sources and then copying them to their own data set (Unwin and Becker, 2002).

STANDARDIZATION OF NUTRIENT VALUES DERIVED FROM NATIONAL FCDBs

The documented nutrient data will be standardized at IARC in terms of units and mode of expression and, as necessary, separated according to analytical method and definition of nutrients. The standardization rules will be incorporated into the import functions of the updated FTI database management software.

EVALUATION OF NUTRIENT VALUES DERIVED FROM NATIONAL FCDBs

The aim of the evaluation is to critically assess and evaluate data derived from national FCDBs according to common criteria. During this process, outdated or outstanding values are detected and will be replaced by better ones: missing values—for single values or entire foods—will be completed with nutrient values from other sources.

The common evaluation guidelines will be based on national database evaluation guidelines, as well as from the newly revised book by Greenfield and Southgate (2002, in prep.). These guidelines define the general evaluation process, quality criteria for imputation, calculation and completion of missing values. For example, the guidelines will define preferred analytical methods per nutrient, order of priority for 'foreign' sources, criteria to choose between estimation through a similar food or calculation, and criteria for selecting the most similar food.

According to these evaluation guidelines, compilers will evaluate nutrient values from their national data set using the updated FTI software. This will enable a critical assessment of their own national data through accessing the documented data of the 10 other FCDBs involved in the project. However, this evaluation is very time consuming and depends upon the resources of national compilers.

CALCULATION OF MISSING NUTRIENT VALUES

Calculation of nutrient values is needed when the food or nutrient value is missing in the national database and if the missing values cannot be estimated through nutrient values of another food. Calculations are by definition only approximations to the real composition of the food but these 'best estimates' are still preferable in epidemiological studies compared to missing values. In general, different kinds of recipe calculations are used providing more or less comparable nutrient values (Powers and Hoover, 1989). Algorithms calculations are in general used to calculate nutrient values of cooked foods, which could also be regarded as single ingredient recipe. Both, the different recipe and algorithm calculation systems in use and the different nutrient retention factors applied in these calculations alter the nutrient intake estimations among countries (NORFOODS, 2002).

To eliminate this source of systematic error in the nutrient intakes, it was decided that within the ENDB project the calculation systems and the retention factors should be common to all countries. Two calculation procedures have been defined by EPIC to calculate missing nutrient values, one based on algorithms for missing cooked foods and the second is based on recipes for multi-ingredient foods. They will be tested and updated according to new information. Missing nutrient values of cooked foods will be calculated based on the nutrient values of the corresponding food in its raw form or in a different cooking method. Specific algorithms are then applied, as well as certain specific coefficients such as 'Vitamin and mineral loss coefficients' (also called 'nutrient retention factors'), 'fat loss coefficients', and 'raw-to-cook weight coefficients' (also called weight yield factors). These coefficients have been collected at IARC from the scientific literature (Møller, 1993; Bergström, 1994; Bognar, 1990), own measurements and calculations based on the British food composition table. The British table was used to verify these coefficients using nutrient values from raw and cooked forms of the same food.

Recipe calculation will be used to estimate the nutrient values of complex foods, which are multi-ingredient foods such as cakes, breads, sauces, or soups. In the ENDB, the concept of the 'yield factor method' (Powers and Hoover, 1989) is used, which is also the base of the recipe calculation used in EPIC-SOFT for mixed recipes (Slimani *et al.*, 1999). In this procedure, each ingredient is treated as a food and all coefficients are applied at the ingredient level, which allows the application of food-specific coefficients to each ingredient of the recipe. When coefficients are applied at recipe level, as in the 'British Method', it is sometimes difficult to know which set of nutrient retention factors should be applied to all ingredients of the recipe because for certain recipes it is difficult to decide on the main ingredient, which determines the recipe category—if cereal, milk, fish, meat, egg or vegetable based (Holland *et al.*, 1991). Most of the recipes for complex foods have been collected by the EPIC collaborators, which in the future will have to be checked and completed.

The EPIC-SOFT data has the specific feature that the fat absorbed during cooking, e.g., frying, is treated as a separate food. This means that the nutrient values of 'fried' foods in the ENDB will need to be expressed excluding the contributions from absorbed fat to avoid accounting for the nutrient values of the absorbed fat twice. Hence, in the ENDB, nutrient values of 'fried foods' will always be estimated by algorithm calculations. If an external user was interested in the nutrient values of the fried food including the contribution of the absorbed fat, e.g., to obtain comparable nutrient values with those in national FCDBs, a recipe calculation would need to be performed with the two ingredients, the absorbed fat and the 'fried' food. The advantage of this practice is that the fatty acid composition of the absorbed fat can be taken into account depending on the type of fat reported by the subject.

COMPILATION OF THE ENDB

The ENDB will be compiled using the data sets incorporated into the updated FTI software. The compilation is carried out in three steps: an automatic pre-compilation based on the food match file, a manual compilation by the compilers during the evaluation step and the final compilation with calculations.

The pre-compilation consists of completing the ENDB with nutrient values according to the exact and similar food matches as indicated in the food match file, without any calculation of missing values through algorithms and recipes. This preliminary compilation allows compilers to conveniently evaluate the nutrient values deriving from the national data sets using the software functions and all other documented national data sets. During the evaluation process, nutrient values will be compiled manually, either through replacing outdated values, completing missing values, or through indicating new food matches according to the evaluation guidelines. The final compilation of the ENDB will consist of incorporating calculated values either through algorithms (missing cooked foods) or through recipe calculation as defined within this project. In every step, control for completeness and consistency will be incorporated into the software.

DISCUSSION AND FUTURE ISSUES

In the past, a lot of work has been done on the harmonization of FCDBs and to produce national FCDBs, reflecting the best possible values of the composition of foods consumed in the country within the limitations of funding, e.g., for additional chemical analysis and new sampling of foods. Correct sampling is important to produce representative nutrient values of a food for a given time and country, which is very costly and time consuming and can therefore only be done for selected foods (Greenfield and Southgate, 1992; Nordbotten *et al.*, 2000; Pehrsson *et al.*, 2000). Despite these limitations, new FCDB editions with revised and updated nutrient values are being published in Europe since 2000 (Germany, the Netherlands, Norway and the U.K.).

Based on this previous work, the ENDB will be the first standardized FCDB with about 100 documented, standardized and evaluated nutrient values for 1000 foods of 10 European countries and a documented data set for the U.S.A. It is a starting point for a more comprehensive and dynamic European FCDB. The ENDB database will not include all foods and components that might be of interest for nutritional epidemiology. It should therefore be extended to cover more foods, components, time periods and countries. However, the ENDB database will allow epidemiological studies to improve the quality of nutrient intake measurements and consequently the results of the investigations of diet and health.

There are also a number of methodological issues that still have to be solved before compositional data can be interchanged internationally. Even though in the ENDB project, agreement was obtained on food classification and description systems, component nomenclature and definition, description of analytical method, calculation methods, retention factors and data imputation procedures, there is a need for global international agreement on these points and their application at national level. In addition, in some countries copyright on food composition data exists and this will have to be dealt with.

Another area of future investment could lie in better collaboration between study designers and food composition experts, especially at the international level, to achieve better data comparability and quality. This means that consumption data should be collected using the same food nomenclature, description and classification and type of recipe handling as used for food composition data. Despite a previous agreement on food classification as the Eurocode (Kohlmeier and Portvliet, 1992), new propositions are being made for an international food classification for dietary surveys (Ireland *et al.*, 2002) because the Eurocode system caused problems in its application and was not used often (Schroll *et al.*, 1996). This collaboration between specialists in food consumption and food composition could improve the availability of compositional data for frequently consumed foods and minimize potential errors in food matching and thus obtain higher precision in estimations of nutrient intakes. In view of the increasing market for functional and novel foods, the collection of compositional data should include foods at product name level.

In the EPIC study, depending on the centre, 43–63% of the food intake is derived from manufactured foods (Charrondiere *et al.*, 2001c). Figure 2 shows the percentage of food intake derived from manufactured foods per country and the number of product names reported in the EPIC study. Commercial products have a wide variation in their composition and parts of them are fortified. For the precision of nutrient intakes, it becomes obvious that in Europe there is a need to obtain more food composition data on a product name level and more compositional data of these products, which should then be included in national FCDBs. The minimum information manufactures should be able to provide are the nutrient values listed on the labels. One has to keep in mind that some nutrient values on labels are derived from food composition tables, either from own or foreign country. A standardized FCDB could therefore improve the quality and comparability of nutrient labels across Europe. Special importance could be given to fortified foods and drinks as they may contribute substantially to micronutrient intakes in some parts of Europe (ILSI Europe Report Series, 1997). Additionally, between 8 and 61% of EPIC subjects in the different centres take vitamin and mineral supplements, with the highest proportion in Middle and Northern Europe (Charrondiere *et al.*, 2001c). This underlines the need of an additional database on vitamin and mineral supplements. The access to compositional data of manufactured foods and vitamin and mineral supplements will, through better nutrient intake estimations, enable a more realistic interpretation of nutrient–health relationship, biomarker measurements, and comparisons with recommended daily intakes (RDIs) and other standards. If the ENDB could be extended to include additives, it could become an excellent tool for risk assessments of chemicals in foods and diet in the area of food safety.

Despite the awareness of food composition experts on the further need for methodological improvements and the development of standardized FCDBs (EFCOSUM group, 2001), researchers in nutritional epidemiology, politicians and

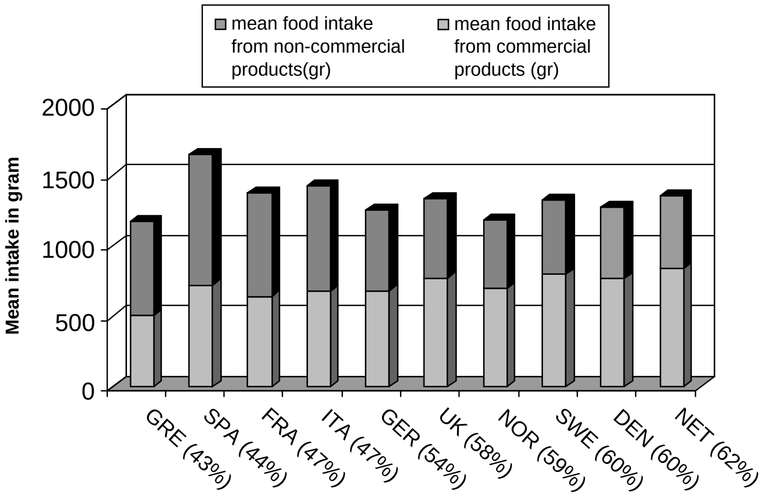

FIGURE 2. Mean food intake in gr (excl. beverages and VMS): from commercial and non-commercial products.

funding agencies are not yet fully aware of the actual problems concerning the incompatibility and imprecision of compositional data and nutrient intake measurements across countries, which could enhance funding for future international methodological work and the compilation of standardized FCDBs.

CONCLUSION

As there is a clear need for a standardized food composition database at European level, the ENDB project was started with 11 countries to compile a documented, standardized and evaluated nutrient database using common rules, formats and calculation procedures. In the future, compositional data from manufactures will be needed for processed foods and vitamin and mineral supplements to improve the nutrient intake estimations, as in general the results obtained using solely existing FCDBs seem to underestimate some nutrient intakes, especially for vitamins and minerals. Outstanding methodological and analytical issues on FCDBs have to be solved before comprehensive and internationally accepted databases can be compiled. There is also a great need to communicate the limitations of actual FCDBs and their impact on nutrient intake estimations to the research community, policy makers and to funding agencies.

REFERENCES

Adelman, M., Dwyer, J. T., Woods, M., Bohn, E., and Otradovec, C. L. (1983). Computerized dietary analysis systems: a comparative view. *J. Am. Diet. Assoc.* **83,** 421–429.

ASEANFOODS (1996). ASEAN food composition tables 2000. Puwastien, P., Burlingame, B., Raroengwichit, M., Sungpuag, P. Institute of Nutrition, Mahidol University, Thailand.

Bergström, L. (1994). *Nutrient losses and gains in the preparation of foods.* EUROFOODS-ENFANT nlg Project nutrient Losses and Gains, No. 32/1994, National Food Administration, Sweden, 1995-08-25.

Bingham, S. A. (1987). The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutr. Abstracts Rev.* **57,** 705–742.

Bognar, A. (1990). *Naehrwerttabellen fuer verzehrte Speisen.* AID Verbraucherdienst, Bonn, Germany.

Cameron, M. E. and Van Staveren, W. A. (1988). *Manual on Methodology for Food Consumption Studies.* Oxford Medical Publications, Oxford.

CEECFOODS (in preparation).

Charrondiere, U. R., Vignat, J., and Riboli, E. (2001a). Differences in fibre intake of a British cohort due to the application of the British, Danish and French food composition tables (FCT). In *Nutrition and Life Style: Opportunities for Cancer Prevention.* Eds. Riboli, E. and Lambert, R. IARC Scientific publications number 156, IARC Press, Lyon 2002.

Charrondiere, U. R., Vignat, J., and Riboli, E. (2001b). Comparable nutrient intake across countries is only possible through standardization of existing food composition tables (FCT). In *Nutrition and Life Style: Opportunities for Cancer Prevention.* Eds. Riboli, E. and Lambert, R. IARC Scientific publications number 156, IARC Press, Lyon 2002.

Charrondiere, U. R., Vignat, J., Slimani, N., and Riboli, E. (2001c). How accurate can nutrient intake be estimated due to the increasing amount of manufactured foods in our diets? In *Proceedings of the Fourth International Food Data Conference*, Bratislava, Slovakia, 24–26 August.

Charrondiere, U. R., Vignat, J., Slimani, N., Riboli, E., and Day, N. (2000). Economic approach to calculate standardized and calibrated nutrient intake across the 10 European Countries involved in EPIC. In *Proceedings of the Fourth International Conference on Dietary Assessment Methods*, in Tucson, U.S.A., 17–20 September 2000.

Deharveng, G., Charrondiere, U. R., Slimani, N., Southgate, D. A. T., and Riboli, E. (1999). Comparison of nutrients in the food composition tables available in the nine European countries participating in EPIC. *Eur. J. Clin. Nutr.* **53,** 60–79.

Eck, L. H., Kledges, R. C., Hanson, C. L., Baranowski, T., and Henske, J. (1988). A comparison of four commonly used nutrient database programs. *J. Am. Diet. Assoc.* **88,** 602–604.

EFCOSUM group. (2001). *European Food Consumption Survey Method.* Final report to the European Commission, Directorate-General Health and Consumer Protection; Directorate G—Public Health. Reference number VS/1999/5182.

Granado, F., Olmedilla, B., Blanco, I., Gil-Martinez, E., and Rojas-Hidalgo, E. (1997). Validation in the intercomparison of food carotenoid content data: user's point of view. *Crit. Rev Food Sci. Nutr.* **37,** 621–633.

Greenfield, H. and Southgate, D. A. T. (1992). *Food Composition Data: Production, Management and Use.* Elsevier, London and New York.

Greenfield, H. and Southgate, D. A. T. (2002). *Food Composition Data: Production, Management and Use*, in preparation.

Guilland, C., Aubert, R., Lhuissier, M., Peres, G., Montagnon, B., Fuchs, F., Merlet, N., and Astorg, P. O. (1993). Computerized analysis of food records: role of coding and food composition database. *Am. J. Clin. Nutr.* **47,** 445–453.

Haraldsdottir, J. (1993). Minimising error in the field: quality controls in dietary surveys. *Eur. J. Clin. Nutr.* **47** (Suppl. 2), S19–S24.

Holland, B., Welch, A. A., Unwin, I. D., Buss, D. H., Paul, A. A., and Southgate, D. A. T. (1991). The Composition of Foods (5th revised and extended edition). Royal Society of Chemistry. Cambridge, U.K.

Hoover, L. W. (1983). Computerized nutrient databases: I. Comparison of nutrient analysis systems. *J. Am. Diet. Assoc.* **82,** 501–505.

ILSI Europe Report Series (1997). *Antioxidants: Scientific Basis, Regulatory Aspects and Industry Perspectives*, 28pp.

Internal report (2001). *Guidelines on how to complete file 9 with the links between national FCDB and EPIC occurrences.* ENDB project.

Ireland, J., van Erp-Baart, A. M. J., Charrondiere, U. R., Møller, A., Smithers, G., and Trichopoulou, A.—for the EFCOSUM group. (2002). Selection of a food classification system and food composition database for future food consumption surveys. *Euro. J. Clin. Nutr.* **56** (Suppl. 2), S33–S45.

Klensin, J. C. (1992). *INFOODS Food Composition Data Interchange Handbook.* The United Nations University, Tokyo, Japan.

Kohlmeier, L. (1991). Problems and pitfalls of food-to-nutrient conversion. In *Their Use in Nutrition Policy-making* (W. Becker and E. Helsing, Eds.), Food and Health Data, pp. 73–84. WHO Regional Publications, European Series No. 34, WHO, Copenhagen.

Kohlmeier, L. and Poortvliet, E. J. (1992). *Eurocode 2 Food Coding System Version 92/1*. Berlin Institute for Social Medicine and Epidemiology.

LATINFOODS (2000). Tabla de composición de Alimentos de América latina.

Lee, R. D., Nieman, D. C., and Rainwater, M. (1995). Comparison of eight microcomputer dietary analysis programs with the USDA nutrient data base for standard reference. *J. Am. Diet. Assoc*. **95,** 858–867.

Lupien, J. R. (1995). The FAO/UNU food composition initiative. *Food Chem*. **57,** 171–173.

Møller, A. (1993). *Loss of Nutrients During Preparation/Cooking: a Practical, Standardized and Systematic Approach*. Report of the Third Annual Meeting of the FLAIR-Eurofoods-Enfant Project in Portugal, 10–12 November 1993.

Mongeau, R. and Brassard, R. (1989). A comparison of three methods for analyzing dietary fiber in 38 foods. *J. Food Comp. Anal*. **2,** 189–199.

Moreiras, O., van Staveren, W. A., Amorim Cruz, J. A., Nes, M., and Lund-Larsen, K.—for the Euronuts SENECA investigators. (1991). Intake of energy and nutrients. *Eur. J. Clin. Nutr*. **45** (suppl. 3), 105–119.

Nieman, D. C., Buterworth, D. E., Nieman, C. N., and Lee, K. E. (1992). Comparison of six microcomputer dietary analysis system with the USDA Nutrient Data Base for Standard Reference. *J. Am. Diet. Assoc*. **92,** 48–56.

Nordbotten, A., Løken, E. B., and Rimestad, A. H. (2000). Sampling of potatoes to determine representative values for nutrient content in a national food composition table. *J. Food Comp. Anal*. **13,** 369–378.

NORFOODS (2002). *Intake Calculation Methods of Food Consumption Data in the Nordic Countries* [in Swedish with English summary]. TemaNord 2002:x, Nordic Council of Ministers, Copenhagen, in press.

Pehrsson, P. R., Haytowitz, D. B., Holden, J. M., Perry, C. R., and Beckler, D. G. (2000). USDA's national food and nutrient analysis program: food sampling. *J. Food Comp. Anal*. **13,** 379–390.

Powers, P. M. and Hoover, L. W. (1989). Calculating the nutrient composition of recipes with computers. *J. Am. Diet. Assoc*. **89,** 224–232.

Riboli, E. (1989). Methodological issues in the investigation of diet and cancer in humans. In *Diet and the Aetiology of Cancer* (A. B. Miller, Ed.), ESO Monograph Series, pp. 55–63. Springer, Heidelberg.

Riboli, E. and Kaaks, R. (1997). The EPIC project: rationale and study design. *Int. J. Epidemiol*. **26** (Suppl. 1), S6–S14.

Schlotke, F., Becker, W., Ireland, J., Møller, A., Ovaskainen, M. L., Monspart, J., and Unwin I. D. (2000). *EUROFOODS Recommendations for Food Composition Database Management and Data Interchange*. European Commission, COST report EUR 19538; http://food.ethz.ch/cost99/datax/.

Schroll, K., Carbajal, A., Decarli, B., Martins, I., Grunenberger, F., Blauw, Y. H., and De Groot, C. P. M. G. (1996). Food patterns of elderly Europeans. *Eur. J. Clin. Nutr*. **50** (suppl. 2), S86–S100.

Shanklin, D., Enders, J. M., and Sawicki, M. (1985). A comparative study of two nutrient data bases. *J. Am. Diet. Assoc*. **85,** 308–313.

Slimani, N., Charrondière, U. R., van Staveren, W. A., and Riboli, E. (2000a). Standardization of food composition databases for the European Prospective Investigation into Cancer and Nutrition (EPIC): general theoretical concept. *J. Food Comp. Anal*. **13,** 567–584.

Slimani, N., Deharveng, G., Charrondiere, U. R., Van Kappel, A. L., Ocke, M. C., Welch, A., Lagiou, A., Van Liere, M., Agudo, A., Pala, V., Brandstetter, B., Andren, C., Stripp, C., Van Staveren, W. A., and Riboli, E. (1999). Structure of the standardized computerized 24-hour diet recall interview used as reference method in the 22 centers participating in the EPIC project. *Comput. Methods Programs Biomed*. **58,** 251–266.

Slimani, N., Ferrari, P., Ocke, M., Welch, A., Boeing, H., van Liere, M., Pala, V., Amiano, P., Lagiou, A., Mattison, I., Stripp, C., Engeset, D., Charrondière, U. R., Buzzard, M., van Staveren, W. A., and Riboli, E. (2000b). Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concept and preliminary results. *Eur. J. Clin. Nutr*. **54,** 900–917.

Slimani, N., Riboli, E., and Greenfield, H. (1995). Food composition data requirements for nutritional epidemiology of cancer and chronic diseases. In *Quality and Accessibility of Food-related Data* (H. Greenfield, Ed.), pp. 209–215. AOAC International, Arlington, U.S.A.

Southgate, D. A. T. (1993). Problems Associated with Different Values for Similar Foods. In *Report of the Third Annual Meeting of the FLAIR-Eurofoods-Enfant Project in Portugal*, 10–12 November 1993.

Unwin, I. D. and Becker, W. (2002). Software management of documented food composition data. *J. Food. Comp. Anal*. **15,** 491–497.

Vandenlangenberg, G. M., Brady, W. E., Nebeling, L. C., Block, G., Forman, M., Bowen, P. E., Stacewicz-Sapuntzakis, M., and Mares-Perlman, J. A. (1996). Influence of using different sources of carotenoid data in epidemiological studies. *J. Am. Diet. Assoc.* **96,** 1271–1275.

Vignat, J., Ireland, J., Møller, A., Unwin, I. D., and Charrondière, U. R. (2001). Guideline notes for preparing and exporting food composition data according to the common formats of export files. *Proceedings of the Fourth International Food Data Conference*, Bratislava, Slovakia, 24–26 August 2001.

Willett, W. (1998). Overview of nutritional epidemiology, and future research directions. In *Nutritional Epidemiology* (W. Willett, ed.), (2nd Edn.) *Monographs in Epidemiology and Biostatistics*, **30,** pp. 1–17, 484–496. Oxford University Press, New York, Oxford.

APPENDIX 1: PARTNERS OF THE ENDB PROJECT

Compilers	International studies	Others
Denmark: Anders Møller *France*: Jayne Ireland *Germany*: Christine Klemm, Andrea Springer *Greece*: Efthimia Grilli, Effi Vasilopoulou *Italy:* Simonetta Salvini *The Netherlands:* Susanne Westenbrink *Sweden*: Wulf Becker *Spain:* Andreu Farran *U.K.*: Susan Church *U.S.A.*: Joanne Holden	*EPIC* *IARC* Elio Riboli Ruth Charrondière Jérôme Vignat Nadia Slimani *EPIC centres* *Denmark*: Connie Stripp *France*: Maryvonne Niravong *The Netherlands*: Marga Ocke, Jose Drijvers, Maryse Niekerk *Norway*: Guri Skeie, Dagrun Engeset *SENECA* Wija van Staveren, The Netherlands *EURALIM* Lluis Serra, Spain	*Information expert* Ian Unwin, U.K. *Food industry* Detlef Mueller, Procter & Gamble, Schwalbach, Germany

APPENDIX 2: LIST OF COMPONENTS TO BE INCLUDED IN THE ENDB

First priority nutrients:

- Factors: density/specific gravity, edible portion, water (moisture), nitrogen conversion factor, fatty acid conversion factor.
- Carbohydrate: total and available carbohydrate, starch, sugars, natural sugar, added sugar, monosaccharides, disaccharides, fructose, galactose, glucose, glycogen, lactose, maltose, maltotriose, ribose, raffinose, stachyose, sucrose (saccharose), xylose.
- Dietary fibre.
- Protein: protein, nitrogen.
- Fat: total fat (lipid), monounsaturated fatty acids, polyunsaturated fatty acids,
- other polyunsaturated fatty acids (= PUFA–linoleic–linolenic), saturated fatty acids, trans fatty acids.
- Alcohol (ethanol).
- Polyols: total polyols, inositol, mannitol, sorbitol, xylitol.
- Organic acids: total organic acids, acetic acid, benzoic acid, quinic acid, citric acid, fumaric acid, di-keto-cholanic acid, iso-citric acid, lactic acid, D-lactic

acid, L-lactic acid, malic acid, oxalic acid, phytic acid (phytin P), propionic acid, salicylic acid. succinic acid, tartaric acid.
- Minerals: ash (minerals), calcium, iron, sodium, non-haem iron, haem iron, potassium, selenium.

Second priority nutrients:

- Dry matter.
- Cholesterol.
- Oligosaccharides.
- Fatty acids: fatty acids 4:0–8:0, fatty acid 10:0 (capric acid), fatty acid 12:0 (lauric acid), fatty acid 14:0 (myristic acid), fatty acid 14:1 (myristoleic acid), fatty acid 15:0 (pentadecylic acid), fatty acid 16:0 (palmitic acid), fatty acid 16:1 (palmitoleic acid), fatty acid 16:1 n-7, fatty acid 16:2, fatty acid 16:4, fatty acid 17:0 (margaric acid), fatty acid 17:1 (heptadecenoic acid), fatty acid 18:0 (stearic acid), fatty acid 18:1 (octadecenoic acid), fatty acid 18:1 n-7, fatty acid 18:1 n-9, fatty acid 18:1 trans n-9 (elaidic acid), fatty acid 18:2, fatty acid 18:2 cis,cis n-6 (linoleic acid), fatty acid 18:3, fatty acid 18:3 n-3 (alpha-linolenic acid), fatty acid 18:4 (stearidonic acid), fatty acid 18:4 n-3 (parinaric acid), fatty acid 20:0 (arachidic acid), fatty acid 20:1 (eicosenoic acid), fatty acid 20:2 (eicosadienoic acid), fatty acid 20:3 (eicosatrienoic acid), fatty acid 20:4 (eicosatetraenoic acid), fatty acid 20:4 n-6 (arachidonic acid), fatty acid 20:5 (eicopentaenoic acid), fatty acid 22:0 (behenic acid), fatty acid 22:1 (docosenoic acid), fatty acid 22:1 n-11 (cetoleic acid), fatty acid cis 22:1 n-9 (erucic acid), fatty acid 22:2 (docosadienoic acid), fatty acid 22:5 (docosapentaenoic acid), fatty acid 22:5 n-3 (clupanodonic acid), fatty acid 22:6 (docosahexaenoic acid), fatty acid 24:0 (lignoceric acid), fatty acid 24:1 (selacholeic acid), fatty acid 4:0 (butyric acid), fatty acid 6:0 (caproic acid), fatty acid 8:0 (caprylic acid).
- Thiamin (vitamin B1)
- Riboflavin (vitamin B2).
- Biotin.
- Vitamin B6, total.
- Folate/vitamin B9: total folate, bound folate, free folate, folic acid.
- Vitamin B12 (cobalamin).
- Vitamin C: vitamin C, L-ascorbic acid.
- Retinol (pre-formed vitamin A).
- Carotenoids: alpha-carotene, beta-carotene, beta-carotene equivalents (provitamin A carotenoids), cryptoxanthin, lutein, lycopene.
- Vitamin E: alpha-tocopherol equiv. from E vitamin activities, alpha-tocopherol, beta-tocopherol, delta-tocopherol, gamma-tocopherol, alpha-tocotrienol, beta-tocotrienol, delta-tocotrienol, gamma-tocotrienol.
- Vitamin D.
- Polyphenols: catechin, coumestrol, daidzein, genistein, isoflavonoids, kaempferol, lignans, quercetin.
- Nitrates.
- Nitrites.